Journal of Environmental Management 166 (2016) 285–293

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

Prototypic automated continuous recreational water quality monitoring of nine Chicago beaches

Dawn A. Shively [a, b], Meredith B. Nevers [a, *], Cathy Breitenbach [c], Mantha S. Phanikumar [b], Kasia Przybyla-Kelly [a, b], Ashley M. Spoljaric [a, b], Richard L. Whitman [a]

[a] U.S. Geological Survey, Great Lakes Science Center, 1574 N County Rd. 300E, Chesterton, IN 46304, USA
[b] Michigan State University, Department of Civil & Environmental Engineering, East Lansing, MI 48824, USA
[c] Chicago Park District, 541 N. Fairbanks, Chicago, IL 60611, USA

A R T I C L E I N F O

Article history:
Received 31 March 2015
Received in revised form
7 October 2015
Accepted 8 October 2015
Available online xxx

Keywords:
Escherichia coli
Beach management
Water quality standards
Predictive empirical modeling
Great Lakes

A B S T R A C T

Predictive empirical modeling is used in many locations worldwide as a rapid, alternative recreational water quality management tool to eliminate delayed notifications associated with traditional fecal indicator bacteria (FIB) culturing (referred to as the persistence model, PM) and to prevent errors in releasing swimming advisories. The goal of this study was to develop a fully automated water quality management system for multiple beaches using predictive empirical models (EM) and state-of-the-art technology. Many recent EMs rely on samples or data collected manually, which adds to analysis time and increases the burden to the beach manager. In this study, data from water quality buoys and weather stations were transmitted through cellular telemetry to a web hosting service. An executable program simultaneously retrieved and aggregated data for regression equations and calculated EM results each morning at 9:30 AM; results were transferred through RSS feed to a website, mapped to each beach, and received by the lifeguards to be posted at the beach. Models were initially developed for five beaches, but by the third year, 21 beaches were managed using refined and validated modeling systems. The adjusted R^2 of the regressions relating *Escherichia coli* to hydrometeorological variables for the EMs were greater than those for the PMs, and ranged from 0.220 to 0.390 (2011) and 0.103 to 0.381 (2012). Validation results in 2013 revealed reduced predictive capabilities; however, three of the originally modeled beaches showed improvement in 2013 compared to 2012. The EMs generally showed higher accuracy and specificity than those of the PMs, and sensitivity was low for both approaches. In 2012 EM accuracy was 70–97%; specificity, 71–100%; and sensitivity, 0–64% and in 2013 accuracy was 68–97%; specificity, 73–100%; and sensitivity 0–36%. Factors that may have affected model capabilities include instrument malfunction, non-point source inputs, and sparse calibration data. The modeling system developed is the most extensive, fully-automated system for recreational water quality developed to date. Key insights for refining and improving large-scale empirical models for beach management have been developed through this multi-year effort.

1. Introduction

Directives for recreational water quality (RWQ) are issued for swimmers' safety, and informing the public as soon as possible of water quality conditions that might threaten their health is a key goal of these documents (Armitage et al., 1999; WHO, 2003; European Parliament, 2006; U.S. EPA, 2012). Predictive, empirical modeling, which estimates fecal indicator bacteria (FIB) concentrations in real-time based on hydrometeorological conditions, has been useful as one of the rapid monitoring tools to meet this goal (Nevers and Whitman, 2005, 2011; Stidson et al., 2012). Use of empirical modeling has increased in recent years (Francy et al., 2013; Thoe and Lee, 2013), and while implementation of EMs at individual beaches has been successful, these EMs have not realized their full potential in scope, functionality, automation, or timeliness of reporting. Many recent EMs rely on samples or data collected manually, which adds to analysis time and increases the burden to the beach manager. Beach-specific instrumentation (hydrological/

* Corresponding author.
E-mail address: mnevers@usgs.gov (M.B. Nevers).

http://dx.doi.org/10.1016/j.jenvman.2015.10.011
0301-4797/© 2015 Published by Elsevier Ltd.

water quality buoys, weather stations), telemetry, and computer programs are useful for improving the performance, rapidity, and ease of model implementation (Francy and Darner, 2007; Francy et al., 2009; Thoe and Lee, 2013). By taking advantage of the available technology, researchers can create fully automated modeling systems for RWQ. Automation permits continuous informing on water quality conditions at the beaches, providing the beach manager with the tools to make informed, immediate decisions about issuing swimming advisories.

Researchers have long recognized the relationship between environmental conditions and concentrations of ambient FIB in beach water (Fujioka et al., 1981; Krogh and Robinson, 1996; Whitman et al., 1999). Some early modeling efforts used rainfall-based preemptive swimming advisories based on the long-understood correlation between FIB and rainfall (Armstrong et al., 1996; Crowther et al., 2001; Lipp et al., 2001; Noble et al., 2003). Results of research by Crowther et al. (2001) showed highly significant increases in FIB after rainfall, and Ackerman and Weisberg (2003) found that rainfall of at least 6 mm warranted preemptive public health warnings. Early EMs that included multiple variables were often developed to explore factors that influenced or explained *Escherichia coli* densities, but few were implemented as beach management tools (Crowther et al., 2001; Nevers and Whitman, 2005, 2010; Nevers et al., 2011). Data for these efforts were usually obtained from the nearest weather station, fixed offshore buoys, or manually generated data from field and laboratory measured variables (e.g., wave height, laboratory measured water chemistry, ambient conditions).

Scale of modeling efforts is limited by local to regional conditions; in some instances, extended shorelines (up to 40 km) exhibit simultaneous fluctuations in FIB concentrations (Whitman and Nevers, 2008), but in other areas, EMs can only be specific to a single beach due to localized physical constraints (e.g., beach angle, contours, bounding structures) or proximity to a point source of contamination (Olyphant and Whitman, 2004; Nevers and Whitman, 2005; Nevers et al., 2007). Further, scale of implementation is driven by equipment requirements and data retrieval.

The use of data loggers and telemetry can improve the breadth and timeliness of EMs by incorporating automatic data retrieval via cell phone or website. Aside from the improvement of hydrometeorological equipment, there has been some research into the best mathematical approach for developing and standardizing EMs (Nevers et al., 2011; Brooks et al., 2013; Mavani et al., 2014). These improvements greatly decrease demands on beach managers.

The purpose of this study was to develop a fully automated water quality management system for multiple beaches managed by the Chicago Park District (CPD) using EMs and state-of-the-art technology (instrumentation, telemetry, and automation) and to compare model effectiveness to the traditional monitoring approach. The modeling system developed is the most extensive, fully-automated system for recreational water quality developed, to date. Further, the model's application likely impacts the largest number of beachgoers of any beach predictive model developed and may serve as a prototype for beach management on an international scale.

2. Methods

2.1. Study site and sample collection

The city of Chicago includes 24 beaches along the southwest shore of Lake Michigan; 21 of which were encompassed in this study. These beaches host upwards of 20 million beach visits during the swimming season (Nevers and Whitman, 2011), which lasts from late May through early September. In this study, we considered nine Chicago, Illinois beaches from north to south, covering approximately 42 km of shoreline: Leone, Osterman, Foster, Montrose, Oak, Ohio, 63rd, Rainbow, and Calumet (Fig. 1). Details and characteristics of the study area can be found elsewhere (Whitman and Nevers, 2008; Nevers and Whitman, 2011). The study period encompassed three years, June 3, 2011 through Sept. 21, 2011; May 25, 2012 through Sept. 10, 2012; and May 23, 2013 through Sept. 12, 2013. *E. coli* data for modeling exercises were obtained from the CPD regular beach monitoring program; beaches were regularly monitored for *E. coli* at least 5 days a week.

2.2. Hydrometeorological data

Water quality buoys were installed at all five of the study beaches in 2011 and at seven in 2012 and 2013. All data buoys (NexSens CB-100) were installed in the swimming area (~1.5 m depth) of the study beaches; data obtained were turbidity (FTS sensor, DTS-12), wave height and wave period (NexSens Accustage pressure transducer; OEM, Keller America), transducer depth, and water temperature.

Multi-parameter weather sensors (Vaisala WXT520) were installed on a light post at three of the study beaches (Foster, Oak, and 63rd); installation heights were 35' (Foster), 20' (Oak), and 30' (63rd). Weather stations were installed in June, 2011 (Foster and 63rd) and on July 1st (Oak). Data obtained were measurements of wind direction and speed, air temperature, rainfall, solar radiation (LI-COR sensor; LI-200), relative humidity, and barometric pressure. Weather stations were left up year-round, only being removed for periodic upgrades or repairs.

Buoy installation took place over a two-week period in June, 2011 at five beaches: Foster, Montrose, Oak, 63rd, and Calumet. Buoys were removed after each swimming season. In 2012, buoy installations took place at seven study beaches between May 14 and July 30 (Leone, Osterman, Montrose, Ohio, 63rd, Rainbow, and Calumet) and in 2013, buoys were installed in April and May.

For all sensors, data were collected hourly from an average of 2 and 5 min (four readings per second) for weather and wave data, respectively; turbidity data were instantaneous. Because there is often a lag between hydrometeorological events and effect on *E. coli* densities (Whitman et al., 2004), hydrometeorological data were aggregated over various time periods prior to the approximate average sample collection time of 10:00 AM. Hourly hydrological data collected from the buoys were aggregated over 4 h, rainfall aggregations were performed over 6, 12, 24, and 48 h, and the remaining weather variables were aggregated over 4 and 6 h. Since Oak's weather station was installed later in the 2011 season and its solar sensor was not functional throughout the season there were insufficient data for use in model building.

2.3. Statistical analysis

Statistical analyses were performed using SPSS version 12.0 (SPSS, 2003) and SYSTAT version 13 (Systat, 2009). Statistical procedures were performed using $\log_{10}$-transformed *E. coli* data to meet parametric assumptions of equality of variance and normal distribution; Kolmogorov–Smirnov test was used to test normality ($P < 0.0001$).

2.4. Empirical model development

Regression analysis using the best sub-set model method in SYSTAT 13 was employed for EM development, and a five-variable maximum was specified to increase parsimony. Criteria considered in determining the best sub-set model were adjusted R square, Akaike information criterion (AIC), AIC corrected, and Bayesian